



Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

ON AN UNSYMMETRICAL PROBABILITY CURVE.

[Second paper.]

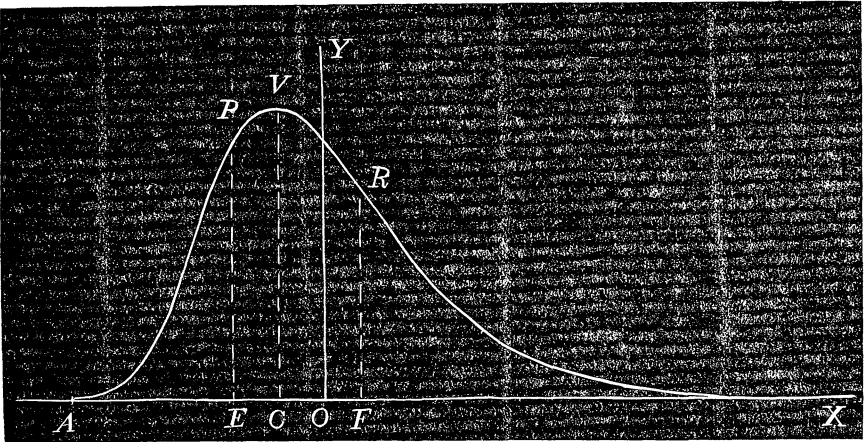
BY E. L. DE FOREST.

IN my article on this subject ending at p. 7 of Vol. X of the ANALYST, some points of interest were omitted or but slightly touched upon. It is proposed to consider them here, and the formulas and tables obtained will be numbered consecutively with the previous ones.

The accompanying figure has been drawn to represent the curve as found for the example at page 1, its ordinates being proportional to those in column 9 of the table. For the equation of any curve of this nature, putting $Ydx = y$, we have by (52),

$$Y = \frac{1}{K \sqrt{(2\pi b)}} \left(1 + \frac{x}{ab}\right)^{a^2b-1} e^{-ax}, \quad (81)$$

where a and b are constants determined by (39) from the q. m. error ϵ and the c. m. inequality ζ , and K is a given function of a^2b .



The curve is in general limited on one side, where it touches the X axis at A , and extends to infinity on the other side, the axis being an asymptote to it in the direction OX . We may say, therefore, that it has a long side and a short side. Since in (81) we have

$$\frac{x}{ab} = \frac{ax}{a^2b},$$

and a^2b is always positive, it appears that Y is a function of ax , and that its value is the same when a and x are both negative, as when they are both positive. Hence, changing the sign of a reverses the curve.

Our figure is drawn for positive a , but if it were negative the curve would be of the same form, only turned over so that the long side would come on the left hand and the short side on the right. From the nature of the curve as shown in (41) and elsewhere, it is evident that if the whole area included between it and the X axis could be cut out of cardboard or other material of uniform thickness, it would balance on a knife edge set to coincide with the Y axis. The greater leverage of the long side compensates for the greater area of the short side. The vertex of the curve does not lie in the Y axis. The maximum ordinate CV is separated from OY by the interval OC , and this is important, for it shows that when the law of error of a set of observations is such that ζ is not zero, the arithmetical mean is not the most probable value of the observed quantity. To get the most probable value of a single observation, we must apply to the arith. mean a correction Δ equal to OC , and by (22) and (39) its amount will be

$$\Delta = -\frac{1}{a} = \frac{-\zeta^3}{2\epsilon^2}. \quad (82)$$

For the observations in Table II. we have

$$\Delta = -\frac{1}{a} = -\frac{1}{1.095} = -0^{\circ}.91,$$

and this added to the arith. mean $5^{\circ}.20$ gives $4^{\circ}.29$ as the value of the amplitude most likely to occur. The probabilities that an error or deviation which occurs will fall on the $+$ or $-$ side of the most probable value, are not the same. The part VCX of the area of the curve lying on the long side is greater than the part ACV lying on the short side, as will appear hereafter. An ordinate dividing the total area into halves, would be between OY and CV . The probable error of the most probable value of an observation is greater toward the long side of the curve than toward the short side. The ordinate which divides the area ACV into halves is PE , and that which divides VCX into halves is RF . The probable error on the $+$ side therefore is CF , and that on the $-$ side is CE .

As found at page 3 and 4, putting $n = a^2b$ and $v = ax$, the areas on the short and long side of any ordinate Y whose abscissa x is reckoned from A as an origin will be respectively

$$\left. \begin{aligned} \int_0^x Y dx &= \frac{v^n e^{-v}}{\Gamma(n+1)} \left(1 + \frac{v}{n+1} + \frac{v^2}{(n+1)(n+2)} + \&c. \right), \\ \int_x^\infty Y dx &= \frac{v^{n-1} e^{-v}}{\Gamma(n)} \left(1 + \frac{n-1}{v} + \frac{(n-1)(n-2)}{v^2} + \&c. \right), \end{aligned} \right\} \quad (83)$$

and their sum is unity. The probability P_2 that an error which occurs will fall on the long side from VC is represented by the area VCX , and its

value is found from the second integral by taking as the lower limit $x = ab - (1 \div a)$, whence we have

$$v = ax = a^2b - 1 = n - 1, \\ P_2 = \frac{(n-1)^{n-1} e^{-n+1}}{\Gamma(n)} \left(2 + \frac{n-2}{n-1} + \frac{(n-2)(n-3)}{(n-1)^2} + \&c. \right), \quad (84)$$

and the series terminates if n is an integer. The values of P_2 have been computed thus, to 7 places of decimals, for $n = 4, 5, 6, 7, 8, 10, 14, 20, 30, 50, 100, 200$. The corresponding values of $P_1 = 1 - P_2$, together with intermediate values found by interpolation, are given to 4 places in the subjoined Table III. By means of the column of differences A_1 , P_1 and consequently P_2 can be found for any given value of the argument, with sufficient accuracy for all practical purposes. Now to get the position of the points

TABLE III.

a^2b	P_1	A_1	$-ar_1$	A_1	ar_2	a^2b	P_1	A_1	$-ar_1$	A_1	ar_2
4	.3528	102	.822	89	1.613	12	.4209	61	1.870	197	2.667
4.5	.3630	82	.911	84	1.703	14	.4270	72	2.067	263	2.865
5	.3712	69	.995	79	1.788	17	.4342	52	2.330	231	3.128
5.5	.3781	59	1.074	76	1.868	20	.4394	69	2.561	359	3.359
6	.3840	52	1.150	71	1.944	25	.4463	45	2.920	333	3.718
6.5	.3892	45	1.221	68	2.016	30	.4508	72	3.253	587	4.051
7	.3937	40	1.289	66	2.085	40	.4580	41	3.840	494	4.639
7.5	.3977	36	1.355	65	2.151	50	.4621	69	4.334	923	5.133
8	.4013	62	1.420	121	2.216	70	.4690	43	5.257	1220	6.056
9	.4075	51	1.541	113	2.337	100	.4733	43	6.477	1321	7.276
10	.4126	45	1.654	110	2.450	140	.4776	36	7.798	1323	8.597
11	.4171	38	1.764	106	2.561	200	.4812		9.121		9.920
					∞	.5000					

E and F which mark the limits of probable error, we must find x from each of the conditions

$$\int_0^\infty Y dx = \frac{1}{2} P_1, \quad \int_x^\infty Y dx = \frac{1}{2} P_2,$$

and by (83) this amounts to solving the numerical equations

$$\left. \begin{aligned} u_1 &= \left(1 + \frac{v}{n+1} + \frac{v^2}{(n+1)(n+2)} + \&c. \right) - \frac{n!}{2} P_1 e^v v^{-n} = 0, \\ u_2 &= \left(1 + \frac{n-1}{v} + \frac{(n-1)(n-2)}{v^2} + \&c. \right) - \frac{(n-1)!}{2} P_2 e^v v^{-n+1} = 0. \end{aligned} \right\} \quad (85)$$

value is found from the second integral by taking as the lower limit $x = ab$ — $(1 \div a)$, whence we have

$$v = ax = a^2b - 1 = n - 1, \\ P_2 = \frac{(n-1)^{n-1} e^{-n+1}}{\Gamma(n)} \left(2 + \frac{n-2}{n-1} + \frac{(n-2)(n-3)}{(n-1)^2} + \&c. \right), \quad (84)$$

and the series terminates if n is an integer. The values of P_2 have been computed thus, to 7 places of decimals, for $n = 4, 5, 6, 7, 8, 10, 14, 20, 30, 50, 100, 200$. The corresponding values of $P_1 = 1 - P_2$, together with intermediate values found by interpolation, are given to 4 places in the subjoined Table III. By means of the column of differences A_1 , P_1 and consequently P_2 can be found for any given value of the argument, with sufficient accuracy for all practical purposes. Now to get the position of the points

TABLE III.

a^2b	P_1	A_1	$-ar_1$	A_1	ar_2	a^2b	P_1	A_1	$-ar_1$	A_1	ar_2
4	.3528	102	.822	89	1.613	12	.4209	61	1.870	197	2.667
4.5	.3630	82	.911	84	1.703	14	.4270	72	2.067	263	2.865
5	.3712	69	.995	79	1.788	17	.4342	52	2.330	231	3.128
5.5	.3781	59	1.074	76	1.868	20	.4394	69	2.561	359	3.359
6	.3840	52	1.150	71	1.944	25	.4463	45	2.920	333	3.718
6.5	.3892	45	1.221	68	2.016	30	.4508	72	3.253	587	4.051
7	.3937	40	1.289	66	2.085	40	.4580	41	3.840	494	4.639
7.5	.3977	36	1.355	65	2.151	50	.4621	69	4.334	923	5.133
8	.4013	62	1.420	121	2.216	70	.4690	43	5.257	1220	6.056
9	.4075	51	1.541	113	2.337	100	.4733	43	6.477	1321	7.276
10	.4126	45	1.654	110	2.450	140	.4776	36	7.798	1323	8.597
11	.4171	38	1.764	106	2.561	200	.4812		9.121		9.920
					∞	.5000					

E and F which mark the limits of probable error, we must find x from each of the conditions

$$\int_0^x Y dx = \frac{1}{2} P_1, \quad \int_x^\infty Y dx = \frac{1}{2} P_2,$$

and by (83) this amounts to solving the numerical equations

$$\left. \begin{aligned} u_1 &= \left(1 + \frac{v}{n+1} + \frac{v^2}{(n+1)(n+2)} + \&c. \right) - \frac{n!}{2} P_1 e^v v^{-n} = 0, \\ u_2 &= \left(1 + \frac{n-1}{v} + \frac{(n-1)(n-2)}{v^2} + \&c. \right) - \frac{(n-1)!}{2} P_2 e^v v^{-n+1} = 0. \end{aligned} \right\} \quad (85)$$

To illustrate the application of the table, take the amplitudes in Table II., for which we have obtained $4^{\circ}.29$ as the most probable value. At p. 3 we find $a^2b = 5.57$. With this argument, Table III. gives

$$\begin{aligned} -ar_1 &= 1.085, & ar_2 &= 1.879; \\ \therefore r_1 &= \frac{-ar_1}{-a} = \frac{1.085}{1.095} = -.99, \\ r_2 &= \frac{ar_2}{a} = \frac{1.879}{1.095} = 1.72. \end{aligned}$$

These are *CE* and *CF* in the figure. It is an even chance whether amplitudes below $4^{\circ}.29$ will be over or under $3^{\circ}.30$, and an even chance whether those above $4^{\circ}.29$ will be over or under $6^{\circ}.01$.

To show what appears to be the proper method of combining observations when the c. m. inequality is of accidental origin, we will now find the most probable mean value and probable error of an angle observed with a theodolite by the method of repetition. The example is taken from Meyer's *Wahrscheinlichkeitsrechnung*, p. 274, where it is treated, of course, in the usual way, by the symmetrical law of probability. Fourteen observations are taken, each giving the angle $17^{\circ} 56'$, plus a number of seconds which stand in our Table IV. in the column headed *u*. The number of repetitions to each observation gives its weight, which stands in column *p*. The next column shows the products *pu*. The general arith. mean is

$$u_0 = \frac{[pu]}{[p]} = \frac{1830}{46} = 39''.78.$$

Subtracting this from each *u*, we get the series of residuals *v*. These are

TABLE IV.

No.	<i>u</i>	<i>p</i>	<i>pu</i>	<i>v</i>	<i>pv</i> ²	<i>p</i> ² <i>v</i> ³
1	45''.00	5	225.0	5.22	136.25	3556.0
2	31 .25	4	125.0	— 8.53	291.04	— 9930.4
3	42 .50	5	212.5	2.72	37.00	503.0
4	45 .00	3	135.0	5.22	81.75	1280.2
5	37 .50	3	112.5	— 2.28	15.60	— 106.6
6	38 .33	3	115.0	— 1.45	6.30	— 27.5
7	27 .50	3	82.5	—12.28	452.40	—16666.2
8	43 .33	3	130.0	3.55	37.80	402.7
9	40 .63	4	162.5	.85	2.88	9.8
10	36 .25	2	72.5	— 3.53	24.92	— 176.0
11	42 .50	3	127.5	2.72	22.20	181.1
12	39 .17	3	117.5	— .61	1.11	— 2.1
13	45 .00	2	90.0	5.22	54.50	569.0
14	40 .83	3	122.5	1.05	3.30	10.4
		46	1830.0		1167.05	— 20396.6

to be reduced to unit weight, before combining them for the purpose of finding the law of error of u_0 . According to (60), the q. m. error of unit weight is $\sqrt{p}$ times that of weight p , and the c. m. inequality of unit weight is $\sqrt[3]{p^2}$ times that of weight p . In other words, in a system of errors whose weight is unity, the mean of the squares of the errors is p times, and the mean of the cubes of the errors taken with their signs is p^2 times what they would be in a system of errors having the weight p . Hence, as is well known, in order to find the approximate q. m. error ε of unit weight from the squared residuals, each v^2 of weight p must be multiplied by p . Likewise, to find the c. m. inequality ζ of unit weight from the cubed residuals, each v^3 of weight p must be multiplied by p^2 . The last two columns in the table show the products pv^2 and p^2v^3 . The number of observations being $m = 14$, we have for the square of the true q. m. error of unit weight

$$\varepsilon_1^2 = \frac{[pv^2]}{m-1} = \frac{1167.05}{13} = 89.77,$$

and as shown at p. 6, by virtue of the relation

$$\zeta_1^3 = \frac{[p^2v^3]}{m} + \frac{\zeta_1^3}{m^2},$$

the cube of the true c. m. inequality of unit weight is

$$\zeta_1^3 = \frac{m[p^2v^3]}{m^2-1} = \frac{14(-20396.6)}{195} = -1464.4. \quad (89)$$

Then for the arith. mean, whose weight is $[p] = 46$,

$$\left. \begin{aligned} \varepsilon_0^2 &= \frac{\varepsilon_1^2}{[p]} = \frac{89.77}{46} = 1.9515, \\ \zeta_0^3 &= \frac{\zeta_1^3}{[p]^2} = \frac{-1464.4}{46^2} = -.6921. \end{aligned} \right\} \quad (90)$$

These are the (q. m. e.)² of u_0 and the (c. m. i.)³ of its possible errors, and from them we can construct the unsymmetrical curve which represents approximately the probabilities of all the possible errors of u_0 .

By (39) we have

$$\left. \begin{aligned} a &= \frac{2\varepsilon_0^2}{\zeta_0^3} = 2\left(\frac{1.9515}{-.6921}\right) = -5.640, \\ b &= \varepsilon_0^2 = 1.9515. \end{aligned} \right\} \quad (91)$$

Since a is negative, the curve has the reversed position. By (82) the most probable error of u_0 here is

$$A_0 = -\frac{1}{a} = \frac{-\zeta_0^3}{2\varepsilon_0^2}, \therefore A_0 = \frac{1}{5.640} = 0''.18. \quad (92)$$

This is more probable than the error zero, or any other. Hence, for the best value of the mean we must take not u_0 alone, but

$$u_0 + A_0 = 39''.78 + 0''.18 = 39''.96.$$

To find the probable errors of this result, we enter Table III. with the argument

$$a^2b = 1.9515(5.640)^2 = 62.1,$$

and there obtain

$$-ar_1 = 4.892, \quad ar_2 = 5.691.$$

The probable errors are therefore

$$r_1 = \frac{-ar_1}{-a} = \frac{4.892}{5.640} = 0''.87,$$

$$r_2 = \frac{ar_2}{a} = \frac{5.691}{5.640} = 1''.01,$$

and the most probable value of the mean is written

$$17^\circ 56' 39''.96 \begin{matrix} +0.87 \\ -1.01 \end{matrix}. \quad (93)$$

We have this final result instead of $17^\circ 56' 39''.78 \pm 0.94$ as obtained by the ordinary method. Table III. also gives the probabilities that the error of u_0 will fall on the short or the long side of the curve from $u_0 + \Delta_0$, meaning in this case error in excess or defect respectively. They are

$$P_1 = .4663, \quad P_2 = 1 - P_1 = .5337.$$

While (93) shows an even chance for the error to fall within the limits $+0.87$ and -1.01 , the chance of its being between 0 and 0.87 is not the same as that of its being between 0 and -1.01 . The probabilities of the two events are

$$\frac{1}{2}P_1 = .2331, \quad \frac{1}{2}P_2 = .2669.$$

We may observe that when the law of probability is unsymmetrical, the ordinary rule that probable error varies inversely as the square root of the weight, will not hold good. The probable errors depend not only on ϵ but on ζ , and these, as (60) shows, vary with the weight by different laws. An observation of weight p is supposed to represent the arith. mean of p observations chosen at random, out of a set of observations of unit weight sufficiently numerous to show the unsymmetrical law of error.

If the c. m. inequality of a set of observations of equal weight, is purely accidental, the probable value of ζ^2 is roughly expressed in terms of ϵ^2 by the relation (80). When observations are of unequal weight, ζ^2 and ϵ^2 varying with the weight in different ratios, we may consider (80) as approximately valid for the average weight of the observations, so that ϵ and ζ denote the q. m. error and c. m. inequality for the weight $[p] \div m$. Then according to (60)

$$\epsilon^2 = \left(\frac{m}{[p]} \right) \epsilon_1^2, \quad \zeta^2 = \left(\frac{m}{[p]} \right)^2 \zeta_1^2.$$

Substituting these in (80) we get

$$(\zeta_1^3) = \pm \frac{.6745 \varepsilon_1^3}{m} \sqrt{(15[p])}. \quad (94)$$

This is the limit within which there is about an even chance for ζ_1^3 to fall if the inequality is fortuitous. In our present example it gives

$$(\zeta_1^3) = \pm \frac{.6745}{14} \sqrt{[15 \times 46(89.77)^3]} = 1076.4.$$

The actual value by (89) is $\zeta_1^3 = -1464.4$. The excess is not very great, and taken in connection with what we know of the nature of theodolite observations, it does not justify the belief that there was any real want of symmetry in the law of error.

Granting that the true law is symmetrical, the most probable value of the observed quantity is the ultimate arith. mean, that is, the mean of all the possible values, each taken a number of times proportional to the probability of its occurrence. But this ultimate mean cannot be found without taking a number of observations which is very large indeed, or infinite. We have to infer its probable value from a moderate number, in the foregoing case only $m = 14$. Owing to the small number of observations, the distribution of the residual errors is irregular and unsymmetrical, and the law of error of the mean, so far as it can be deduced from the observations, is an unsymmetrical law. Hence, the most probable value of the ultimate arith. mean is not the apparent mean u_0 , but u_0 plus the correction Δ_0 due to the c. m. inequality ζ_0 . The resulting value of the mean and its probable errors as exemplified in (93) constitute, as it seems to me, the best determination we can make of the most probable value of the observed quantity.

CORRESPONDENCE.

Editor Analyst:

I desire to make an additional remark in regard to the paradox in the Query by Professor Johnson. (See ANALYST, No. 2, p. 44.)

The principle which seems to be violated in the Ex. given is this:—If $u = f(a, x), \dots (1)$ and $du \div dx = f'(a, x), \dots (2)$ and we assign to a such value as will reduce u to a constant or to zero, then the same value ought to reduce $f'(a, x)$ to zero. If it does not, the paradox must arise from a misinterpretation of the result, or of some of the *singular forms* that enter into the work.

Now although $\sin \infty$ and $\cos \infty$ may each be regarded as zero, in the sense that any constant may be regarded as zero in comparison with another quantity which is infinitely greater, we cannot admit 0 as an absolute value in violation of the principle $\sin^2 \infty + \cos^2 \infty = 1$.